MULTIVARIATE ANALYSIS ASSIGNMENT 1

TEAM INFORMATION

Ruixin Yang (RUID:)

Aishwarya Senthilvel (RUID: 199001269)

DESCRIPTION

The project we are going to analyze is about the effect of a marketing campaign of a commercial bank in Portugal. This campaign is about promoting the sales of the banks' term deposits. We want to know what influences client's decisions about buying term deposits. In this analysis, the factor we are looking at is what key factor captures client's decisions. And there are bunch of other factors affecting this decision from various respects. To specify these factors, we hereby propose our hypothesis below.

Hypothesis 1: We assume that clients' characteristics is relevant to the decision of buying term deposits.

Hypothesis 2: We assume that clients' previous contacts are relevant to the decision of buying term deposits.

Hypothesis 3: We assume that the performance of markets is relevant to the decision of buying term deposits.

SPAP ANALYSIS

Based on these hypothesizes above, we start presenting our SPAP analysis.

1. S.M.A.R.T Goal:

   Determine clients' decisions of buying term deposits from a bank

2. Measures of the dependent variable:

   I. The binary variable that shows whether a client subscribes to a term deposit. (Data.world database, binary values summed by dates) Based on which, we calculate the ratio of 1 (yes to the term deposit).

   II. The binary variable that shows whether a client subscribe a term deposit. (Data.world database, binary values summed by names of clients) Based on which, we calculate the ratio of 1 (yes to the term deposit).

   III. The binary variable that shows whether a client subscribe a term deposit. (Data.world database, binary values summed by names of deposit products) Based on which, we calculate the ratio of 1 (yes to the term deposit).

3. Propose potential aspects for independent variables:

      I.   Do clients' characteristics influence their decisions of buying term deposits?

     II.   Do clients' previous contacts influence their decisions of buying term deposits?

   III.   Does the performance of market influence their decisions of buying term deposits?

4. Determine independent variables:

      I.   From hypothesis 1 and the first question above, we can see that independent variables may contain age, type of job, marital status, default record, ongoing housing loan, ongoing personal loan, and education level.

     II.   From hypothesis 2 and the second question above, we can see that independent variables may contain number of days that passed by after the client was last contacted from a previous campaign, contact communication type, last contact month of year, last contact duration, number of contacts both before and during the campaign, and outcome of previous campaign on clients.

   III.   From hypothesis 3 and the third question above, we can see that independent variables may contain employment variation rate, consumer price index, consumer confidence index, EUIRBOR 3-month rate, and number of employees.

5. Evaluate the accessibility of the data:

We know the location: consumer price index, consumer confidence index, EUIRBOR 3-month rate, last contact month of year, number of days that passed by after the client was last contacted from a previous campaign, number of contacts both before and during the campaign, age, type of job, marital status, default record, ongoing housing loan, ongoing personal loan

We only know the existence: number of employees, contact communication type, outcome of previous campaign on clients,

The existence is unclear: employment variation rate, last contact duration, education level

6. For independent variables that the locations are known, here are the table of primitive analysis:

independent variable	chart type	x-axis	y-axis
consumer price index	line	consumer price index	ratios of 1 (Yes) from decisions
consumer confidence index	line	consumer confidence index	ratios of 1 (Yes) from decisions
EUIRBOR 3	line	EUIRBOR 3	ratios of 1 (Yes) from decisions
last contact month of year	bar	last contact month of year	ratios of 1 (Yes) from decisions

Number of days that passed by after the client was last contacted from a previous campaign	line	Number of days that passed by after the client was last contacted from a previous campaign	ratios of 1 (Yes) from decisions
Number of contacts both before and during the campaign	bar	Number of contacts both before and during the campaign	ratios of 1 (Yes) from decisions
age	bar	age	ratios of 1 (Yes) from decisions
type of job	bar	type of job	ratios of 1 (Yes) from decisions
marital status	bar	marital status	ratios of 1 (Yes) from decisions
default record	bar	number of defaults	ratios of 1 (Yes) from decisions
ongoing housing loan	bar	number of housing loans	ratios of 1 (Yes) from decisions
ongoing personal loan	bar	number of personal loans	ratios of 1 (Yes) from decisions

Data Dictionary

TABLE	COLUMN	DATA TYPE	REQUIRED	SENSITIVE	VALUES	DESCRIPTION
Bank	Age	Numeric				
Bank	Job	Categorical			Admin Blue-Collar Entrepreneur Housemaid Management Retired Self-Employed Services Student Technician Unemployed Unknown	Type of job
Bank	Marital	Categorical			Divorced Married Single Unknown	Divorced also covers widowed
Bank	Education	Categorical			Primary Secondary Tertiary	

09/17/2020

Bank	Default	Categorical			No Yes Unknown	Has credit in default?
Bank	Balance					
Bank	Housing	Categorical			No Yes Unknown	Has housing loan?
Bank	Loan	Categorical			No Yes Unknown	Has personal loan?
Bank	Contact	Categorical			Cellular Telephone	Contact Communication Type
Bank	Day	Categorical			Mon Tue Wed Thu Fri	Last contact day of the week
Bank	Month	Categorical			Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec	Last contact month of year
Bank	Duration	Numeric				Last contact duration in seconds
Bank	Campaign	Numeric				Number of contacts performed during this campaign and for this client

Bank	Pdays	Numeric				Number of days that passed by after the client was last contacted from a previous campaign * 999 means client was not previously contacted
Bank	Previous	Numeric				Number of contacts performed before this campaign and for this client
Bank	Poutcome	Categorical			Failure Nonexistent Success	Outcome of the previous marketing campaign
Bank	y	Binary			Yes No	Has the client subscribed a term deposit
Bank Additional	Emp_var_rate					Quarterly indicator pf employment variation rate
Bank Additional	Cons_price_idx					Consumer price index - monthly indicator
Bank Additional	Cons_conf_idx					Consumer confidence index - monthly indicator
Bank Additional	euribor3m					Euribor 3 month rate - daily indicator
Bank Additional	Nr_employed					Number of employees - quarterly indicator